

ALIBABA CLOUD

# 阿里云

## 数据湖分析

### MongoDB

文档版本：20201013

 阿里云

## 法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

# 通用约定

格式	说明	样例
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告 重启操作将导致业务中断，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击设置> 网络> 设置网络类型。
<b>粗体</b>	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
<code>Courier</code> 字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
<i>斜体</i>	表示参数、变量。	<code>bae log list --instanceid</code> <i>Instance_ID</i>
[ ] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

# 目录

1. 读写MongoDB数据	05
1.1. 创建MongoDB表并读写数据	05

# 1.读写MongoDB数据

## 1.1. 创建MongoDB表并读写数据

本文主要介绍如何创建MongoDB表并读。

### 创建MongoDB表

MongoDB Schema创建成功之后，接下来就可以创建MongoDB表。在DMS中通过以下示例在MongoDB中创建person表：

```
create external table dla_person (  
  id int,  
  title varchar(127),  
  age int,  
  create_time timestamp  
)TBLPROPERTIES (  
  TABLE_MAPPING = 'person',  
  COLUMN_MAPPING = 'title,name';  
);
```

- external: 创建的表是外表
- tblproperties: 外表与源表的映射关系。
  - table\_mapping: 表名映射
  - column\_mapping: 字段映射关系

表创建成功后，您可以通过MySQL客户端或者MySQL命令行工具连接DLA中的MongoDB Schema，从而使用SQL语句读写MongoDB数据库。

### 通过DLA将OSS中的数据写入MongoDB

对存储在OSS/Table Store上的大数据进行分析之后，可以通过DLA把结果数据回写到MongoDB，供前台业务使用。

以dla\_person表为例，通过MySQL命令行工具连接DLA，执行以下SQL语句把oss\_db中customer的前十条记录进行一些转换，然后插入了mongo\_test.dla\_person表：

```
mysql> insert into mongo_test.dla_person
  -> select c_custkey, c_name, c_custkey + 20, now() from oss_db.customer limit 10;
+-----+
| rows |
+-----+
|  10 |
+-----+
1 row in set (3.72 sec)
mysql> select * from mongo_test.dla_person;
+-----+-----+-----+-----+
| id | title      | age | create_time      |
+-----+-----+-----+-----+
|  1 | james      |  10 | 2018-12-14 14:22:54.369 |
|  2 | bond       |  20 | 2018-12-14 14:23:48.527 |
|  3 | lily       |  30 | 2018-12-14 14:23:48.962 |
|  4 | lucy       |  20 | 2018-12-14 14:23:49.396 |
|  1 | Customer#000000001 | 21 | 2018-12-20 10:15:56.629 |
|  3 | Customer#000000003 | 23 | 2018-12-20 10:15:56.629 |
|  5 | Customer#000000005 | 25 | 2018-12-20 10:15:56.629 |
|  7 | Customer#000000007 | 27 | 2018-12-20 10:15:56.629 |
|  9 | Customer#000000009 | 29 | 2018-12-20 10:15:56.629 |
|  2 | Customer#000000002 | 22 | 2018-12-20 10:15:56.629 |
|  4 | Customer#000000004 | 24 | 2018-12-20 10:15:56.629 |
|  6 | Customer#000000006 | 26 | 2018-12-20 10:15:56.629 |
|  8 | Customer#000000008 | 28 | 2018-12-20 10:15:56.629 |
| 10 | Customer#000000010 | 30 | 2018-12-20 10:15:56.629 |
+-----+-----+-----+-----+
14 rows in set (0.16 sec)
```

### 查询MongoDB嵌套字段表

假设MongoDB中存在嵌套数据结构的collection, details 中 city 字段是一个嵌套的字段。

```
db.Ha.insert({
  id: 1,
  name: "james",
  age: 1,
  details: {
    city: "hangzhou"
  }
});
```

DLA中可通过以下SQL创建表：

```
CREATE EXTERNAL TABLE Ha (
  id int,
  name string,
  age int,
  city string
)
TBLPROPERTIES (
  COLUMN_MAPPING = 'city,details.city;'
)
```

**COLUMN\_MAPPING**：将DLA中声明的字段映射到底层MongoDB字段。

DLA中查询表数据：

```
mysql> select * from Ha;
+-----+-----+-----+-----+
| id | name | age | city |
+-----+-----+-----+-----+
| 1 | james | 1 | hangzhou |
```

## 查询ObjectId类型的字段

如果想要查询MongoDB中的ObjectId字段，例如\_id。

DLA中可通过以下SQL创建表：

```
CREATE EXTERNAL TABLE Ha (  
  _id ObjectId,  
  id int,  
  name string,  
  age int,  
  city string  
);
```

DLA中查询表数据：

```
mysql> select * from Ha where _id = ObjectId('5edef42407bcd42c8bad4284');  
+-----+-----+-----+-----+-----+  
|_id          | id | name | age | city |  
+-----+-----+-----+-----+-----+  
| 5e de f4 24 07 bc d4 2c 8b ad 42 84 | 1 | james | 1 | hangzhou |
```